Readability approaches: Implications for Turkey

Mustafa Ulusoy
University of Illinois at Urbana-Champaign *ulusoy@uiuc.edu*

Finding the right fit between students' reading ability and textbooks is very important for comprehension. Readability studies aim to analyse texts to find the right fit between students and texts. In this literature review, readability studies are classified under quantitative, qualitative and combined quantitative-qualitative readability approaches. The quantitative approach includes readability formulas, cloze test, and checklists and scales. The qualitative approach consists of leveling and checklists. The combined qualitative and quantitative readability approach is new in the field. In this approach, readability formulas can be used together with benchmark passages and checklists. The literature shows that readability formulas rely heavily on surface features of a text, and gives a rough estimate of the text readability. The qualitative approach focuses on the quality of writing style, and is criticised as being too subjective. The paper concludes by evaluating the implications of readability studies for Turkey.

Readability, quantitative and qualitative approaches, Turkey, cloze procedure

READABILITY

Readability is an important issue and finding the right fit between students' reading ability and text difficulty is an important and challenging task for teachers (Armbruster, in press; Fry, 1977a). Students have different prior experiences. In addition, every classroom has students who are above and below the average reading level. Reading texts also have a difficulty range. Although some texts can be read easily by the students, others are hard to read. Finding the right fit between the difficulty level of the text and the students' reading ability is very critical.

In this literature review, different readability approaches and research results are investigated. At the end of the review, readability and its possible implications are evaluated for Turkey. Armbruster (in press), and Gunning (2003) classified the major readability approaches as quantitative approaches, qualitative approaches, and combined quantitative/qualitative assessments. This classification is used to organise the review.

Quantitative Approaches

"The beginning of readability research come from two main sources- studies of vocabulary control and studies of readability measurement" (Chall, 1988, p. 2). In the vocabulary control studies, new words in the texts, repetition of these new words, and their difficulty were studied. From this point of view, the vocabulary control studies and the readability studies had similar purposes to measure the difficulty of texts for learning. Readability studies started in the 1920s for two reasons. First, Thorndike's *Teacher's Word Book* enabled teachers and researchers to measure word difficulty objectively. Second, more students started to go to secondary schools in the 1920s. These students did not have strong backgrounds to read the textbooks written for the earlier secondary school population (Chall, 1988).

As readability formulas, cloze tests, and checklists give quantitative results for text evaluators, they can be examined under the quantitative approaches.

Readability Formulas

Readability is usually determined according to a mathematical formula. Syntactic (sentence) and semantic (vocabulary) dimensions are used to apply the readability formulas. Sentence length and number of syllables per sentence are counted to have a number that represent the readability level of the texts. Generally, it is true that shorter sentences are easier to read than longer sentences and short words are easier to read than long words (Jacobson, 1998).

Generally, readability formulas give a rough estimate of text readability. According to Readence, Bean and Baldwin (2001), over 30 different readability formulas and graphs have been developed. On the other hand, Fry (2002) believes that there are over 100 readability formulas in existence.

Fry and SMOG readability formulas are the most popular for teachers to evaluate middle school and junior/senior high texts (Ruddell, 2005). These formulas are not complex. Teachers can use the formula manually because the calculation are very easy. For the Fry readability formula, teachers select three passages. Each should have 100 words. Proper nouns, initialisation, and numerals are counted. Then, the number of sentences and the number of syllables are counted, and converted to an average. Teachers find the intersect point of these two lines on the graph and the graph gives an approximate grade level from 1 to 17+ grades (Fry, 1968, 1977b).

In the SMOG formula, evaluators count 10 consecutive sentences near the beginning of the text, 10 in the middle, and 10 near the end. Every word of three or more syllables is counted in these 30 sentences. Then, the square root of the number of polysyllabic words is counted. Three is added to the approximate square root to find the SMOG grade (Ruddell, 2005).

In addition to hand-calculated formulas, there are other much complex formulas available on software for computer applications. These are Flesch-Kincaid, Degrees of Reading Power (DRP), ATOS (Advantage-TASA Open Standard), and The Lexile Framework. Scanning the whole text is the biggest advantage in the computer based readability formulas. DRP uses the "Dale List" and average number of letters per word; ATOS uses number of words, and average grade level of words, and Lexile Framework uses the sentence length and word frequency (Gunning, 2003). Flesch-Kincaid is a syllable formula, and it is used in Microsoft Word. The readability statistics include counts of words, characters, paragraphs, sentences, averages of sentences per paragraph, words per sentences, and characters per word. Also, the results give the "Flesch Reading Ease" and "Flesch-Kinkaid Grade Level". Microsoft Word users can find the readability statistics under the "Edit" option if they check "Show Readability Statistics" for MAC (Anderson, 2003). The same specialty can be found under the "Tools" option for the Microsoft Word for PCs.

Readability has the strength of objectivity and consistency. Any person or computer gets the same score when they search for the readability of a particular text. In addition, there is very large research base about the validation of readability formulas (Fry, 2002). On the other hand, these formulas rely heavily on surface features such as word and sentence length, and ignore the conceptual load and deeper syntactic structures (Valdes, Barrera, and Cardenas, 1984). Bailin and Grafstein (2001) also believe that students can understand some long sentences better than short ones because sentence length may actually facilitate comprehension, rather than impede it. In addition, the authors claim that vocabulary changes rapidly and different socio-cultural groups have different core vocabularies. Using the same word lists may not be suitable for these different cultural groups.

Readability formulas are used to evaluate literature, social studies, and science texts. Researchers generally examined the readability levels of the materials. Also, the readability literature contains

some other representative research about the validation of the readability formulas for other languages. Some selected research is given in the following section of the review.

Laughlin (1973) determined the readability level of the 79 easy-to-read books receiving primary level grading in the 1971 Children's Catalog through the use of three readability formulas (Fry, Spache, and Wheeler-Smith). The Fry Formula showed that 73 per cent of the books were suitable for the first grade readers. The Spache and Wheeler-Smith formulas revealed that 14 per cent and 86 per cent respectively were readable at the second grade level. The researcher concluded that easy-to-read failed to designate a specific readability level, even within one series.

Another study was conducted by Baxter (1992) to determine the effect of text difficulty as determined by the Fry readability formula on the comprehension monitoring performance of above and below average fourth grade readers whose selection was based on their standardised test scores. In this study, the subjects were given passages at levels that were matched to their reading levels. The hypothesis that their monitoring performance would get worse as the material got harder was not supported. In this study, readability of the material was determined solely by the Fry readability formula and it did not seem to have any pattern of influence on the performance of any of the readers. The researcher concluded that something in the passages or within the readers did cause performance to vary from passage to passage. Also, the researcher claimed that the match between the reader and the text is much more complicated than the simple matching of reader ability as determined by a standardised test with text readability as determined by a readability formula.

Dohrman (1972) investigated the readability of 75 intermediate social studies topic-articles from eight sets of encyclopaedias with two (Dale-Chall and Fry) readability techniques and determined the suitability of individual encyclopaedia articles for intermediate grade use from the resultant predicted readability scores. She evaluated a total of 600 topic-articles from the eight "Best Quality" encyclopaedias recommended by the American Library Association. The results indicated that there was no significant difference among the eight selected encyclopaedias as to their suitability such as collateral readability with students' graded reading levels, for grades four, five and six with regard to the readability of 75 social studies topics evaluated.

Crawford (1984) attempted to develop a valid and reliable procedure to assess the readability of elementary level materials in Spanish. She used average sentence length and average number of syllables per 100 words as independent variables, and grade levels as the dependent variable. By using multiple regression analysis, she developed a formula and a graph that use the same procedure for the Fry readability graph in English.

Parker and Hasbrouck (2001) used two formulas, modified Fry graph and computerised Spaulding formula, developed for Spanish language text to analyse nine stories that were read by 36 Spanish-speaking second graders with Limited English Proficiency (LEP). The nine passages about animals were translated into Spanish by a team of three native Spanish speakers. To find the modified Fry readability score the evaluators counted the number of syllables and sentences from three 100-word Spanish passages and subtracted the number 67 from the average number of syllables before plotting the results on the graph. The Spaulding formula uses average sentence and a list of most frequently used words in Spanish. Students read the stories and experienced evaluators recorded the errors such as substitutions, omissions, mispronunciations, and so on. Researchers found that the modified Fry formula only moderately predicted oral reading accuracy and oral reading fluency. Spaulding's Spanish readability scale was used only to predict chance levels. The researchers concluded that unique words, word length, and density of syllables were important both in English and in Spanish.

Hamsik (1984) conducted a study to determine if four widely used readability formulas (the Flesch formula, the Dale-Chall formula, the Fry Graph, and the Lorge formula) measured reading

difficulty for ESL students enrolled in intensive English centers in the United States in preparation for academic work. The research results showed that the four readability formulas and graphs can be used to measure readability of ESL students and select material appropriate to the reading level of ESL students.

Borchers (1990) asked teachers whether they used 11 reading strategies including the Fry readability graph and cloze tests. Of the 11 reading strategies, readability formulas had the lowest frequency of use for social studies, science and English teachers in their classrooms. This result shows that readability formulas are not popular for content area teachers.

Cloze Procedure

Cloze procedure was introduced by Taylor in 1953 and is based on the person's ability to complete incomplete words, images or thoughts (Ruddell, 2005; Vacca and Vacca, 2005).

> This techniques is used to determine the readability of written material, an individual's reading level on specific material, an individual's vocabulary level in a specific subject or topic area, an individual's language skills, and an estimate of an individual's general comprehension level. (Mariotti and Homan, 2001, p. 137)

Teachers first select a passage of 275-300 words and type the passage in double-space. The first and last sentences are left intact. Then, every fifth word of the other sentences are deleted. If the fifth word is a proper noun, it is skipped and the next word is deleted. Words are eliminated until there are a total of 50 deletions. Evaluators determine the correct replacements and multiply them by 2. Only the correct exact replacements are accepted as correct answer. According to Rankin and Culhane (1969), scores of 60 per cent and higher indicate that passage can be read independently by students. Scores between 40 and 60 per cent mean that students can read the passage with instruction. A score below the 40 per cent indicates that the passage is too difficult for students.

The cloze test can give evaluators more information than readability formulas because this test estimates how well each student functions when they interact with the text. In spite of these advantages, cloze procedure has one big disadvantage. Students generally do not like to do cloze tests because they are difficult for them. Evaluators should provide students five to ten practice sentences before administering the test. They should not expect a valid score when they first administer the test (Mariotti and Homan, 2001; Ruddell, 2005; Vacca and Vacca, 2005).

Checklists and Scales

Researchers developed alternative approaches to seek interactions between readers and texts. Students bring their prior knowledge, interests, and attitudes to these texts. Several instruments were developed to deepen evaluators' knowledge about the texts and textbooks. Irwin and Davis's (1980) "Readability Checklist" and Singer's (1992) "Friendly Text Evaluation Scale" are the best known scales that give quantitative results to the evaluators.

The Readability Checklist provides information about the readability of the text. It also gives teachers information to make a text more readable (Irvin, and Davis, 1980). The checklist has two variables. These are understandability and learnability. Understandability provides information to the teachers about required background knowledge to comprehend the text, and syntactic difficulty of the text such as sentence construction, main idea, and detail arrangements. Learnability gives information about the clarity and usefulness of the text or book (Ruddell, 2005).

Teachers rate the text by using a five point scale from unacceptable to excellent. After completing the checklist, teachers sum up their ratings, determine lowest rated items, and summarise the weaknesses of the text.

The Friendly Text Evaluation Scale is another way to see what makes a text friendlier than another. It is generally used along with the Fry graph. The scale has five variables: (1) text organisation (discourse consistency, and cohesiveness), (2) explication of ideas (prior knowledge, and organisational basis), (3) conceptual density (number of new ideas and vocabulary), (4) meta-discourse (conversation between the author and the reader about the text), and (5) instructional devices (table of contents, headings, subheadings, glossary and index) (Singer, 1992). Evaluators use a five point scale from strongly disagree to strongly agree. Textbook evaluators rate the scale and sum up their ratings at the end. "A score closer to 34 implies the text is friendly; scores closer to 170 suggest the text is unfriendly" (Singer, 1992, p. 163).

Hookstra (1990) used Singer's scale to evaluate three grade six social studies textbooks. Undergraduate and post-baccalaureate students majoring in elementary education performed the evaluation of the selected textbooks. According to the research results, textbooks were moderately considerate in terms of the five areas (text organisation, explication of ideas, conceptual density, meta-discourse, instructional devices). Organisation and instructional devices were rated highest and appropriate level of conceptual density and appropriate use of meta-discourse features were rated lowest. The researcher concluded that the textbooks were readable but the sub-areas of the Singer scale could of course be improved.

Qualitative Approaches

Readability formulas are useful to analyse average number of polysyllabic words and the average number of sentences taken from sample passages in a text. According to Readance, Bean, and Baldwin (2001), these measures are not applicable to poetry, or the symbolic discourse of such disciplines as mathematics, and chemistry.

The quality of the writing style is also an important factor to evaluate in analysing the readability of texts. Readability formulas are not sensible for scrambled passages. They can give the same score for a scrambled passage and for a passage with acceptable syntax (Readance, Bean, and Baldwin, 2001). In addition, some important text variables such as structure, coherence and cohesion, and important reader variables such as prior knowledge, interest, motivation and purpose for reading cannot be measured by readability formulas (Armbruster, in press). Evaluators can use checklists to have ideas about these important text and reader variables. Leveled books can also be used to establish text difficulty.

Checklists

Instead of a checklist, Alvermann and Phelps (2002, p. 174) give teachers a framework to develop their own checklist. The framework has four headings. These are content, format, utility and style. Every heading has its own open ended questions. For example, the content includes questions about the depth of the content, new or difficult vocabulary, new concepts, and appropriateness of the text and students' prior knowledge. Format has questions about illustrations, introductions, summaries and index. Utility includes questions about activities, teacher's manual and additional readings. The last part of the framework is style and it has questions about the complexity and cohesion of the text or book.

Gunning (1998) claimed that teachers should foster students' ability to choose books by helping them develop criteria. Students should be able to select interesting but not difficult books. According to Fielding (as cited in Gunning, 1998, p. 29), students should be able to ask themselves the following questions in order to find appropriate-level books:

1. Does the book seem interesting?
2. Does the book seem to be like the ones I usually read? Does it have about the same number of words? Does it have about the same number of pictures?
3. Can I read most of the words or figure out most of the ones I don't know?
4. When I try reading a page or two, does my reading seem smooth or do I have to stop a lot to try to figure out words?

Armbruster and Anderson (1981) formulated a set of maxims. They claimed that if authors follow the maxims they can produce "considerate" text. These maxims are:

1. Structure: Choose a discourse structure that best conveys the informative purpose.
2. Coherence: Make the relationships among ideas clear enough so that there is a logical connection or flow of meaning from one idea to the next.
3. Unity: Address one purpose at a time; do not stray from the purpose by including irrelevant and distracting information.
4. Audience Appropriateness: Make sure the text fits the knowledge base of the reader (pp. 2-3).

Readers can gather appropriate information from considerate text with minimal cognitive effort. On the other hand, as the maxims are not presented in the inconsiderate text, readers need to spend extra time to apply them for the text. Readers can comprehend an inconsiderate text but they will need to spend extra time, effort, strategy and skill (Anderson and Armbruster, 1984).

Armbruster and Anderson (1981; 1984) developed a "Textbook Evaluation Response Form". Evaluators can use this checklist to evaluate text quality. The authors implied that the checklist is not an objective method to evaluating texts. The checklist can help evaluators to systematise their subjective judgments about text quality. The checklist can be completed in eight steps. First, evaluators select three 150 to 400 words passages from the textbooks. These passages should have a title or heading. Second, evaluators only read the title and formulate questions based on information found in the heading or title. When they prepare the questions, they should remember that these questions will be answered in the remaining of the passage. Third, they read the passage and underline the parts that are necessary to answer the questions. Fourth, they decide if the passages have complete or adequate answers to relevant questions. Fifth, they rate each paragraph by using 1 to 5 scale (1 is Low, and 5 is High). Sixth, evaluators determine how coherent the passage is. They should answer five sub questions to decide about the passage. These questions are related to connectives, pronouns, nouns, phrases, and order of the events that are found in the passage. Seventh, they determine to what extent the ideas in the passage contribute to a single text unit or frame. Finally, evaluator decides how appropriate the text is for readers.

According to Armbruster (in press), checklists have two major weaknesses. First, teachers do not use the checklists because they do not evaluate the textbooks before their adoption. Second, they do not have time and effort to evaluate the texts.

Garner, Slater, Weaver, and Cole (2001) were interested in finding out if pre-service teachers notice inconsiderate text features. They used the characteristics of inconsiderate text identified by Armbruster and Anderson (1981). The pre-service teachers surveyed the intermediate social studies textbooks. These textbooks had the characteristics of the inconsiderate text. Student-teachers were told to read the chapters three times. Then, they tried to find where the students might experience difficulty understanding the material. The results showed that pre-service teachers were not able to detect inconsiderate text features. Only two of the 17 teachers mentioned any flaws, and detected only a single problem of five appearing in the text segment. The researchers concluded that pre-service methods courses need to provide experience for prospective teachers in noticing and compensating text flaws. Also, they concluded that a

checklist developed by Anderson and Armbruster (1984) can guide evaluators in making judgments about textbooks.

Leveling

"Leveling systems are especially important at the beginning levels of reading where type size, number of words on a page, and helpfulness of illustrations can make a significant difference" (Gunning, 2003, p. 180). Leveling takes a number of "text support" factors into consideration such as illustration, length, curriculum, language structure, judgment, and format (Fry, 2002).

There are different kinds of leveling systems. For example, teachers in New Zealand created a readability system where they categorised the books into nine levels, such as emergent, early 1, early 2, early 3, early 4, and fluency 1, fluency 2, fluency 3 and fluency 4 (Gunning, 2003).

Gunning (1998) used another leveling system and provided annotated listing of best books for beginning readers from Picture level to Grade 2A. He believes that "Informal Reading Inventory" is widely accepted placement device for beginning readers. This inventory has students begin with easy material and read increasingly difficult material. As a result of the inventory, teachers can find the right level for the students. He gives a "Graduated Word Lists" and "Primary Reading Passages Inventory" in his book. According to him, they can give reasonably valid placement information. After students have been assigned a level, teachers check the accuracy of the placement by observing the students.

Fountas and Pinnel (1999) also provided leveled books for K-3. They worked with teachers, and developed levels A through P. There is no rigid division between grade levels in their system. According to their approximate grade level correspondence, Kindergarten level is between A-C, Grade 1 is between B-I, Grade 2 is between H-M, Grade 3 is between L-P, and Grade 4 is between O-R levels.

Pinnel and Fountas (2002) gave a leveled books list in their book for grades 3-6. They create a gradient of text. In their gradient, "Level A is Kindergarten, Level B-I is Grade One, Level H-M is Grade Two, Level L-P is Grade Three, Level O-T is Grade Four, Level S-W is Grade Five, Level V-Y is Grade Six, Level Z is Grades Seven, Eight" (p.13).

Another leveling system was created by Rog and Burton (2002). They created explicit descriptions of what texts look like at each level (Level 1-10) of development. These studies show that there are different kinds of leveling systems. As Brabham and Villaume (2002) said, leveled texts are important because they encourage teachers to select materials that are just right for students. On the other hand, leveling has some limitations. First, leveling is used mostly from K-6. Second, the leveling system is qualitative, and it is more subjective (Armbruster, in press).

Combined Quantitative and Qualitative Analyses

"Combining qualitative and quantitative measures of readability is a new trend but it is still in its infancy" (Armbruster, in press). In this trend, a readability formula can be used as a beginning estimate. Then, teachers can judge the text based on a checklist and benchmark passages (Rothkopf, 1985).

The Qualitative Assessment of Text Difficulty (Chall, Bissex, Conard, and Harris-Sharples, 1996) uses both qualitative and quantitative factors. It can be used for both informational and fictional books, and has six scales and nine levels from grade one through college. First, evaluators select an appropriate scale. Then, they can select 100-word examples from their textbook or text. Evaluators compare the selected passage with benchmark passage, and try to determine sentence complexity, conceptual difficulty and vocabulary difficulty, and number of ideas. If teachers

evaluate a book, they can select different 100-word passages and take an average from the selected passages to determine the difficulty.

Kinder, Bursuck and Epstein (1992) combined quantitative and qualitative methods to measure textbook readability. They evaluated 10 eighth-grade American history textbooks. The authors used an evaluation form based on Armbruster's (1984) conception of global and local coherence. Organisational signals and frequently used text structures were used to evaluate global coherence. Local coherence was evaluated on the paragraph level. Pronouns and their antecedent nouns were found. In addition to these, readabilities were calculated using the Fry Readability Scale for each text. The research results showed that all the textbooks were one or more years above the grade level. Texts had most of the elements of organisation but their qualities were different. Although most of the texts had instructions, only 3 per cent provided review material contained in previous chapters and only 60 per cent provided an overview of material to be presented in the chapter. The researchers identified many questions both within and at the end of text chapters but questions at the beginning were almost nonexistent. The local coherence of the text was found good by the researchers.

Readability in Turkey

According to Oliveira (1996), textbooks are more important to developing countries than they are in developed countries. As a developing country, the importance of the textbooks in Turkey supports Oliveria's claim. Textbooks are considered an essential tool for schooling in Turkey. They are highly valued and used by the Ministry of National Education and teachers because textbooks help them deliver the curriculum.

Readability formulas that are prepared for the English speaking world are not suitable for Turkish since the calculated readability levels of first and second grade level textbooks give results at the 17+ grade readability level. Apparently, this is because of the structure of the Turkish language. Leveling system for English books is also not a suitable way to evaluate the readability level of Turkish books.

The best-known readability formulas such as Fry and SMOG should be validated for Turkish. Developing a leveling system for Turkish story books could also be useful for the teachers and students. It is clear that readability studies for Turkish texts and textbooks are waiting for Turkish researchers. Other qualitative and quantitative text readability methods such as checklists, and cloze procedure can be used in their present form to evaluate Turkish textbooks' difficulty level.

Evaluating the Turkish content area textbooks should be the first step for researchers. After evaluating textbook readabilities qualitatively and quantitatively, suggestions could be made to publishers. Students' achievement on cloze test passages can also be an indicator about textbook readability. These kinds of studies would give publishers information about the match between students' reading and textbooks' difficulty levels.

The best match between the students and textbook can be made by teachers. Turkish content area teachers should be able to evaluate readability of textbooks and students' reading abilities. In-service education for content area teachers and reading courses for pre-service content area teachers may be the solution to inform them about these topics. After having a basic knowledge about the readability, these teachers can evaluate the textbooks for their students.

Teachers' evaluations can give publishers more realistic feedback about the match between students' reading and textbooks' difficulty levels. There is no doubt that readability studies for Turkish content area textbooks and storybooks will help publishers to provide students more readable and comprehensible books.

REFERENCES

Alvermann, D. E, and Phelps, S. F. (2002). *Content Reading and Literacy: Succeeding Intoday's Diverse Classroom*. Boston, MA: Allyn and Bacon.

Anderson, M. A. (2003). Everyday tech tools and reading. *Multimedia Schools*, 10(6), 21-23.

Anderson, T. H., and Armbruster, B. B. (1984). Content area textbooks. In R. C. Anderson, J. Osborn, and R. J. Tierney (Eds.), *Learning to Read in American Schools: Basal Readers and Content Texts* (pp. 193-226). Hillsdale, NJ: Lawrence Erlbaum Associates.

Armbruster, B. B. (in press). Matching readers and texts: The continuing quest. In D. Lapp, J. Flood, and N. Farnan (Eds.), *Content area reading and learning* (3rd ed., pp.). New York: Erlbaum.

Armbruster, B. B., and Anderson, T. H. (1981). *Content Area Textbooks* (Reading Education Report No. 23). Urbana, IL: Center for the Study of Reading.

Bailin, A., and Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language and Communication*, 21, 285-301.

Baxter, K. M. (1992). Reader-text match: The interactive effect of reader ability and text difficulty on comprehension monitoring. Unpublished doctoral dissertation, University of Northern Iowa.

Borchers, C. A. (1990). Content area teachers' uses of textbooks and content area reading strategies in Kansas science, social studies, and English secondary school classes. Unpublished doctoral dissertation, Kansas State University.

Brabham, E. G., and Villaume, S. K. (2002). Leveled text: The good news and the bad news. *The Reading Teacher*, 55, 438-441.

Chall, J. S. (1988). The beginning years. In B. L. Zakaluk, and S. J. Samuels (Eds.), *Readability: Its Past, Present, and Future* (pp. 2-13). Newark, DE: International Reading Association.

Chall, J. S., Bissex, G. L., Conard, S. S., and Harris-Sharples, S. H. (1996). *Qualitative Assessment of Text Difficulty: A Practical Guide for Teachers and Writers*. Cambridge, MA: Brookline Books.

Crawford, A. N. (1984). *A Spanish Language Fry-Type Readability Procedure: Elementary Level* (Vol. 7 No. 3). Los Angeles, CA: California State University, Evaluation, Dissemination and Assessment Center.

Dohrman, M. H. (1972). A readability evaluation of selected intermediate grade social studies topics from eight encyclopaedias. Unpublished doctoral dissertation, The University of New Mexico.

Fountas, I. C., and Pinnell, G. S. (1999). *Matching Books to Readers: Using Leveled Books in Guided Reading, K-3*. Portsmouth, NH: Heinemann.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11, 513-516.

Fry, E. (1977a). *Elementary Reading Instruction*. New York: McGraw-Hill.

Fry, E. (1977b). Fry's readability graph: Clarifications, validity, and extension to level17. *Journal of Reading*, 21, 242-252.

Fry, E. (2002). Readability versus leveling. *The Reading Teacher*, 56, 286-291.

Garner, R., Slater, W. H., Weaver V. P., and Cole, M. (2001). Do pre-service teachers notice textbook flaws? *Education*, 106(4), 429-433.

Gunning, T. G. (1998). *Best Books for Beginning Readers*. Needham Heights, MA: Allyn and Bacon.

Gunning, T. G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders*, 23, 175-189.

Hamsik, M. J. (1984). Reading, readability, and the ESL reader. Unpublished doctoral dissertation, The Florida State University.

Hookstra, G. M. (1990). Middle school social studies: An examination of textbook structure, classroom interaction, and student achievement. Unpublished doctoral dissertation, Oregon State University.

Irwin, J. W., and Davis, C. A. (1980). Assessing readability: The checklist approach. *Journal of Reading*, 24, 124-130.

Jacobson, J. M. (1998). *Content Area Reading: Integration with the Language Arts*. New York: Delmar.

Kinder, D., Bursuck, B., and Epstein, M. (1992). An evaluation of history textbooks. *The Journal of Special Education*, 25, 472-491.

Laughlin, M. K. (1973). The readability of the Easy-to-Read trade books. Unpublished doctoral dissertation, The University of Oklahoma.

Mariotti, A. S., and Homan, S. P. (2001). *Linking Reading Assessment to Instruction: An Application Worktext for Elementary Classroom Teachers* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Oliveira, J. (1996). Textbooks in developing countries. In V. Greaney, (ed.) *Promoting Reading in Developing Countries* (pp. 78-90). Newark, DE: International Reading Association.

Parker, R. I., and Hasbrouck J. E. (2001). Spanish readability formulas for elementary-level texts: A validation study. *Reading and Writing Quarterly*, 17, 307-322.

Pinnell, G. S., and Fountas, I. C. (2002). *Leveled Books for Readers Grades 3-6*. Portsmouth, NH: Heinemann.

Rankin, E. F., and Culhane, J. W. (1969). Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 13, 193-198.

Readence, J. E., Bean, T. W., and Baldwin, R. S. (2001). *Content Area Literacy: An Integrated Approach* (7th ed.). Dubuque, IA: Kendall/Hunt.

Rog, L. J., and Burton, W. (2002). Matching texts and readers: Leveling early reading materials for assessment and instruction. *The Reading Teacher*, 55, 348-356.

Rothkopf, E. Z. (1985). Writing to teach and reading to learn: A perspective on the psychology of written instruction. In H. Singer, and R. R. Ruddell (Eds.) *Theoretical Models and Processes of Reading* (3rd ed. pp.879-907). Newark, DE: International Reading Association.

Ruddell, M. R. (2005). *Teaching Content Reading and Writing* (4th ed.). Hoboken, NJ: John Wiley and Sons.

Singer, H. (1992). Friendly texts: Description and criteria. In E. K. Dishner, T. W. Bean, J. E. Readence, and D. W. Moore (Eds.), *Reading in the Content Areas: Improving Classroom Instruction* (3rd ed., pp. 155-170). Dubuque, IA: Kendall/Hunt.

Vacca, R. T., and Vacca, J. A. L. (2005). *Content AREA Reading: Literacy and Learning Across the Curriculum* (8th ed.). Boston, MA: Pearson Education.

Valdes, G., Barrera, R., and Cardenas, M. (1984). Constructing matching texts in two languages: The application of propositional analysis. *The Journal for the National Association of Bilingual Education*, 9, 3-19.